

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic acid search using SW model

Run on: January 14 2003 11:42:33 Search time 37.817 seconds
(without alignments)
11136.303 Million cell updates/sec

Title: US-09-910-428-2
Perfect score: 26
Sequence: 1 cctcccaatcaatcatttcttc 26

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: em_estba:*
2: em_estba:*
3: em_estba:*
4: em_estba:*
5: em_estba:*
6: em_estba:*
7: em_estba:*
8: em_estba:*
9: em_estba:*
10: em_estba:*
11: em_estba:*
12: em_estba:*
13: em_estba:*
14: em_estba:*
15: em_estba:*
16: em_estba:*
17: em_estba:*
18: em_estba:*
19: em_estba:*
20: em_estba:*
21: em_estba:*
22: em_estba:*
23: em_estba:*
24: em_estba:*
25: em_estba:*
26: em_estba:*
27: em_estba:*

Pred. No. is the number of results predicted by Chari's law, a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21.4	82.3	431	17	BH032631
2	21.4	82.3	431	17	BH032631
3	20.6	80.0	473	17	A473609
4	20.6	79.2	973	17	CNS0700N
5	20.4	78.5	790	17	CNS0700N
6	20.2	77.7	804	17	A773248

Result No.	Score	Query Match	Length	DB ID	Description
7	19.8	76.2	475	19	AM522881
8	19.8	76.2	535	10	AM523946
9	19.8	76.2	577	10	BH032631
10	19.8	76.2	587	10	BH032631
11	19.8	76.2	635	17	A2736913
12	19.8	76.2	639	17	A2736913
13	19.8	76.2	644	10	AM331729
14	19.8	76.2	670	17	A7873765
15	19.8	76.2	683	14	B0199536
16	19.8	76.2	702	17	A7350728
17	19.8	75.4	297	17	A0351507
18	19.8	75.4	375	14	F32327
19	19.6	75.4	853	14	B0481609
20	19.6	75.4	1055	17	CNS04YMG
21	19.4	74.6	278	9	A1157727
22	19.4	74.6	364	9	A223404
23	19.4	74.6	421	9	AA200990
24	19.4	74.6	514	17	A2946332
25	19.4	74.6	548	9	A1645945
26	19.4	74.6	826	11	AK012255
27	19.4	74.6	1015	12	BF160577
28	19.2	73.8	220	17	A2595053
29	19.2	73.8	451	17	A0321193
30	19.2	73.8	618	17	A2845764
31	19.2	73.8	638	14	B0597590
32	19.2	73.8	666	17	BH439410
33	19.2	73.8	771	12	BF16123
34	19.2	73.8	718	12	BF17059
35	19.2	73.8	750	17	CNS00977
36	19.2	73.8	808	12	BE869444
37	19.2	73.8	855	12	B6192496
38	18.8	72.3	401	17	A2660399
39	18.8	72.3	470	13	B1418635
40	18.8	72.3	504	17	A0209240
41	18.8	72.3	518	12	B6522866
42	18.8	72.3	526	13	B1417940
43	18.8	72.3	576	17	A2350346
44	18.8	72.3	600	17	CNS02MD5
45	18.8	72.3	862	12	BQ246571

ALIGNMENTS

RESULT 1
LOCUS: BH032631
DEFINITION: PCT-24-239N18-TV PCT-24 Mus musculus genome clone PCT-24-239N18
ACCESSION: BH032631
VERSION: G1:14803637
KEYWORDS: GSS.
SOURCE: house mouse.
ORGANISM: Mus musculus.
REFERENCE: Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akint, R., Levins, M., Russell, D., de Jong, P., and Fraser, C. M. Mouse BAC End Sequences from Library PCT-24 Unpublished (1999)
Other GSS: PCT-24-239N18-TV
COMMENT: Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: sblaw@igrg.org
Clones are deposited from the mouse BAC library PCT-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.nih.gov). Clones may be purchased from BACMAN.

/note="Vector. pMD42uv. Purified genomic DNA from *M. musculus* C57Bl/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/tomorrow.htm>). The DNA was hydrolytically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adapter oligonucleotides were ligated to the blunt ends in high molar excess. The adapter DNA was purified and size selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (911732141b; AF12072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adapters complementary to the insert adapters and purified. The sheared, adaptered mouse DNA was annealed to adapterly competent *E. coli* XL10 Gold (Stratagene) cells and selected for ampicillin resistance."

HASH COUNT 181 a 118 c 115 g 159 t

ORIGIN

Query Match 80 08 Score 20.8 PR 17 Length 973
Best Local Similarity 91.78; Prod. No. 1.5e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Caps 0;

DB 2 CTGCCAATCATATTCTTCT 25
|||||
244 CTGCCAATCATATTCTTCT 267

RESULT 4
CNS00009/c
LOCUS
DEFINITION
973 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TE13 end of BAC:
HACR38C06 of RPCL-98 library from *Drosophila melanogaster* (fruit fly), genomic survey sequence.

ACCESSION
AL076629
AL076629.1 GI:4956106

VERSION
GSS.

KEYWORDS
Drosophila melanogaster.
Drosophila melanogaster.
Pukarijola, M.3-03, Arthropoda, Insecta, Euryptera,
Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha,
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 973)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope, Centre National de Sequencage, BP 191 91006 Evry cedex - FRANCE (E-mail: seqref@genoscope.cns.fr - Web: www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org/The-BDGP/Drosophila-melanogaster-BAC-library>. The BAC library was prepared by Kazuo Ogasawara and Aaron Mammoser in Peter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCL-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp. the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..973
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR38C06"
/clone_lib="RPCL-98"
/note="End : TE13"

HASH COUNT 299 a 152 c 187 g 305 t 30 others

ORIGIN

Query Match 73.28 Score 20.6, EB 17, Length 773;
Best Local Similarity 90.08; Prod. No. 2.2e+02;
Matches 20; Conservative 3; Mismatches 2; Indels 0; Caps 0;

DB 1 CCTGCCAATCATATTCTTCT 25
|||||
973 MCTGCCAATCATATTCTTCT 949

RESULT 5
CNS0700N/c
LOCUS
DEFINITION
770 bp DNA linear GSS 08-JUL-2001
T7 end of clone AZ00A011E07 of library AZ00A from strain CBS 712 of *Kluyveromyces marxianus*, genomic survey sequence.

ACCESSION
AL423957
AL423957.1 GI:12207151

VERSION
GSS.

KEYWORDS
Kluyveromyces marxianus.
Kluyveromyces marxianus.
Eubaryota, Fungi, Ascomycota, Saccharomycotina, Saccharomycetes,
Saccharomycetales, Saccharomycetaceae, Kluyveromyces.
Soucier, J., Aigle, M., Artiguenave, F., Blandin, G.,
Boulotin-Fukuhara, M., Bon, E., Brolier, P., Casaroglia, S.,
de Montigny, J., Duion, B., Durrens, P., Lepinque, A., Llorente, B.,
Malpertuy, A., Neveuglise, C., Ozier, K., Papadopoulos, O., Potter, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Jouvel, M.,
Winkler, P., and Weissenbach, J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of
yeast species for molecular evolution studies
FEMS Lett. 487 (1), 3-12 (2000)

REFERENCE
11152876
2 (bases 1 to 790)
Llorente, B., Malpertuy, A., Blandin, G., Artiguenave, F., Winkler, P.,
and Duion, B.
Genomic exploration of the hemiascomycetous yeasts: 12.
Kluyveromyces marxianus var. marxianus
FEMS Lett. 487 (1), 71-75 (2000)

JOURNAL
MEDLINE
20584722
11152887
3 (bases 1 to 790)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope, Centre National de Sequencage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail: seqref@genoscope.cns.fr - Web: www.genoscope.cns.fr)
This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *avaram*, *Saccharomyces eubayanus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces fragilis* var. *fragilis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia pastoris*, *Debaryomyces hansenii* var. *hansenii*, *Pichia stipitiphila*, *Candida lipolytica* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
Location/Qualifiers
1..790
/organism="Kluyveromyces marxianus"
/strain="CBS 712"
/variety="marxianus"
/db_xref="taxon:4911"
/clone="AZ00A011E07"
/clone_lib="AZ00A"
/note="End : 17"

BASE COUNT 236 a 126 c 155 g 263 t 10 others

ORIGIN

Query Match 78.58; Score 20.4; DB 17; Length 790;

VERSION AWS29446.1 GI:7166341
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 (Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus).

REFERENCE 1 (bases 1 to 535)
 AUTHORS Honaldo,M.F., Lennon,G. and Soares,M.H.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Heksteden Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mssoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NciI site and the oligo-dT track served to identify it as a clone from the normalized hippocampus library cDNA library preparation: M.B. Soares lab clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1 50,
 >AT-richLow-complexity
 Seq primer: M13 forward
 POLYA-Yes.

FEATURES
 source location/Qualifiers
 1..535
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="01-R-B00-ale-q-11-0-01"
 /clone_lib="01-R-B00"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTZ19D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The library (01-R-B00) is a subtracted library derived from a mixture of the following tissues: thalamus, cerebellum, hypothalamus, medulla, pons, midbrain, cerebral cortex, corpus striatum and hippocampus. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.uiowa.edu. The subtraction has been previously described in (Honaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
 TAG_LIB=01-R-B00
 TAG_TISSUE=hippocampus
 TAG_SEQ=CATTC*

BASE COUNT 196 a 107 c 62 g 175 t
 ORIGIN

Query Match 76.2%; Score 19.8; DB 10; Length 535;
 Best Local Similarity 91.3%; Pred. No. 3 aa+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTCGCCAATCAATTCATTTC 24
 |||||
 DB 330 CTCGCCAATCAATTCATTTC 352

RESULT 9 577 bp mRNA linear EST 13-JUN-2000
 LOCUS BE109206
 DEFINITION U1-R-BS1-azd-b-07-0-01.s1 U1-R-BS1 Rattus norvegicus cDNA clone
 SOURCE U1-R-BS1-azd-b-07-0-01 3', mRNA sequence.
 ACCESSION BE109206

VERSION BE109206.1 GI:8501311
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 (Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus).

REFERENCE 1 (bases 1 to 577)
 AUTHORS Honaldo,M.F., Lennon,G. and Soares,M.H.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Heksteden Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: mssoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NciI site and the oligo-dT track served to verify it as a clone from the normalized embryo at 13 dpc library cDNA library preparation: M.B. Soares lab clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 1-50,
 >AT-richLow-complexity
 Seq primer: M13 forward
 POLYA-Yes.

FEATURES
 source location/Qualifiers
 1..577
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="01-R-BS1-azd-b-07-0-01"
 /clone_lib="01-R-BS1"
 /dev_stage="embryonic 13 dpc"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTZ19D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; The 01-R-BS1 library is derived from 13 dpc whole embryo tissue. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.uiowa.edu.
 TAG_LIB=01-R-BS1
 TAG_TISSUE=embryo at 13 dpc
 TAG_SEQ=AATTC*

BASE COUNT 203 a 113 c 77 g 184 t
 ORIGIN

Query Match 76.2%; Score 19.8; DB 10; Length 577;
 Best Local Similarity 91.3%; Pred. No. 4a+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CTCGCCAATCAATTCATTTC 24
 |||||
 DB 330 CTCGCCAATCAATTCATTTC 352

RESULT 10 587 bp DNA linear GSS 18-JUL-2001
 LOCUS RH093286
 DEFINITION RPCI-24-297H24-TV RPCI-24 Mus musculus genomic clone RPCI-24-297H24
 DNA sequence.
 ACCESSION RH093286
 KEYWORDS RH093286.1 GI:14913191
 SOURCE GSS.
 ORGANISM Mus musculus
 (Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

BASE COUNT 241 a 143 c 108 g 188 t
ORIGIN

and selected for ampicillin resistance."

Query Match
Best local similarity 76.2% Score 19.8; DB 17; Length 670;
91.8% Pred. No. 4,400.02;
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

0Y 4 CTGCGCAATCAATACATTTC 25
|||||
DB 493 CTGCGCAATCAATACATTTC 515

RESULT 15

BO199536

18705 683 bp mRNA linear EST 02-MAY-2002

DEFINITION U1-R-ER0 c1c-b-08 0 01 s1 U1-R-ER0 Rattus norvegicus cDNA clone

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.

U1-R-ER0 c1c-b-08 0 01 3', mRNA sequence.